

# **EXHIBIT J**

## EXHIBIT 1

### **Preservation Questions Regarding The New York Times's Sumo Logic Data Repository**

1. The Times has represented that (i) the volume of available data on Sumo Logic is 1.44 GB; and (ii) there are 65,704 individual requests stored in Sumo Logic. Please confirm these figures are accurate, and if not, please provide updated figures.

**Answer:** There were 65,704 individual requests and 1.44 GB of available data stored in Sumo Logic as of March 14, 2025. As of June 27, 2025, there are 82,149 individual requests and 1.97 GB of available data stored in Sumo Logic.

2. Are each of the 65,704 individual requests a single user prompt and response, a single session containing potentially several prompts and responses, or something else?

**Answer:** Each request is a single user prompt and response.

3. Besides ChatExplorer, are there other interfaces or methods used by Times employees, in their business capacity, to call on OpenAI's GPT models, whether via an API or via products such as ChatGPT or Copilot?

**Answer:** Yes.

4. Is there any database or repository besides Sumo Logic of ChatExplorer logs reflecting information regarding Times employees' use of the OpenAI models at issue, the OpenAI products at issue, or the Microsoft products at issue?

**Answer:** Yes. When ChatExplorer logs are stored in Sumo Logic, they are also written to Amazon S3 for long-term storage. As a result, ChatExplorer logs exist in both Sumo Logic and Amazon S3, but the retention policies for the platforms are different. Furthermore, logins to ChatExplorer are recorded in Azure authentication logs. However, Azure authentication logs do not capture user prompts or responses.

5. Please identify the data fields included in the Sumo Logic data and explain the information reflected by each field.

- stream: Text referencing how the log was collected (e.g. stdout)
- timestamp: Timestamp for when the entry occurred (e.g. 1751982642653)
- log.emailAddress: Email address of the user on ChatExplorer (e.g. jane.doe@nytimes.com)
- log.model: Text referencing the model that was used in the API call (e.g. gpt-4o)
- log.user: Text that the user submitted (e.g. fix this)
- log.assistant: Text that the API returned (e.g. Hello there,\n\nThank you so much for reaching out about the \)

6. Please identify the number of individual user IDs in the Sumo Logic data.

**Answer:** 512.

7. Please describe the nature of the Sumo Logic data and how it is stored (e.g., in a local database or cloud-based storage), including all file types and extensions (e.g., JSON/CSV), whether any data fields are encrypted/redacted, and whether the data is in a structured or unstructured format.

**Answer:** Sumo Logic is a cloud-based logging platform. It is agnostic as to the type of data that is sent to it. It has the ability to redact certain fields; however, redactions must be applied to each individual field. There have been no redactions applied to any ChatExplorer logs.

8. Does the Sumo Logic data contain the full contents of the user prompts and LLM responses? If not, where is the other data stored and maintained?

**Answer:** Yes.

9. Does the Sumo Logic data indicate when system or intermediary prompts (e.g., system message, guardrails) are injected automatically before the model is invoked?

**Answer:** No.

10. When did The Times implement a retention policy for Sumo Logic data?

**Answer:** The Times applied a retention policy for Sumo Logic data since the platform was implemented. In March 2024, following OpenAI's assertion that ChatExplorer logs are relevant to this litigation, IT was explicitly instructed by in-house counsel to preserve ChatExplorer logs. However, prior to June 20, 2024, ChatExplorer data was stored in a different Sumo partition that inadvertently had a 90-day retention period. When this was discovered on June 20, 2024, the retention policy on the partition was immediately changed to 180 days to prevent further data loss. ChatExplorer data was then migrated to its own partition with a 360-day retention policy and was simultaneously written to Amazon S3 with a 10-year retention policy. And, existing ChatExplorer data in the old partition was exported and uploaded to Amazon S3.

11. Have any Sumo Logic prompts been deleted since The Times began considering litigation against OpenAI?

**Answer:** Yes. ChatExplorer logs from December 5, 2023, to March 20, 2024, were inadvertently deleted.

12. How many Sumo Logic prompts have been deleted since The Times began considering litigation against OpenAI?

**Answer:** We are unable to determine how many prompts have been inadvertently deleted.

13. Are any Sumo Logic prompts that have been deleted since The Times began considering litigation against OpenAI recoverable?

**Answer:** No.

14. Please identify The Times's internal departments, teams, or groups that use ChatExplorer (e.g., GenAI News Core, Machine Learning Platforms team, Research & Development, Newsroom, XFun, Legal, etc.).

**Answer:** Data and Insights, Digital Products and Technology, International, Marketing, Newsroom, NYT Audio, NYT Licensing, Opinion, Product and Design, Project Management, Shared Services Center, Standalone Products, Technology, Wirecutter, Audience Insights Group, Cybersecurity, Enterprise Tech, and XFun.

15. Please describe or produce your data retention policies for ChatExplorer log data. Who has access to view, search, or export the log data, and what audit log exists of that data?

**Answer:** The Times currently retains ChatExplorer log data in Sumo Logic for 360 days. The logs are also retained in Amazon S3 for 10 years. The majority of users can view ChatExplorer logs stored in Sumo Logic. However, those users are unable to manually delete any of the data.

16. If an employee departs The Times, are the data logs in Sumo Logic associated with that employee subject to The Times's regular data retention policies for former employees?

**Answer:** Yes.